

Appendix A

Technical Notes: Occupational Injury and Illness Incidence Rates and Closed Claims Study

Occupational Injury and Illness Incidence Rates

BLS Survey of Occupational Injuries and Illnesses: The Bureau of Labor and Statistics (BLS), with the help of the state agencies, selects a non-proportional stratified probability sample of employment establishments and mails them questionnaires. Employers are instructed to record all nonfatal employee injury and illness incidents, number of days away from work for each recorded injury/illness, the number of employee hours worked and the establishment's average employment. Participants in the annual survey consist of employers who maintain Occupational Safety and Health Administration (OSHA) records on employee injuries and illnesses on a regular basis under federal law and smaller employers who are exempt from OSHA record keeping requirements. The data collection process differs for the former and the latter. The former are mailed a questionnaire in February, following the survey year, and are asked to transfer from their records all injuries and illnesses incurred as well as demographic and hours worked data. The latter, exempt employers (those with fewer than 11 employees and those designated as "low-hazard industries" by OSHA) are notified in December of the prior year (e.g., contacted in December of 2000 to record injuries for the 2001 survey) that they have been chosen to participate in the survey and must keep records of all employee injuries. The participating state agencies are responsible for collecting data from employers within their jurisdiction and for submitting these questionnaires to BLS for analysis. The BLS uses its incidence rates as a benchmark by which to compare the frequency of injuries and illnesses occurring within jurisdictions, industries or specific occupations for a calendar year. The variable "Total Injuries and Illnesses per 100 Full-time workers" (the most widely quoted measure) is calculated as follows:

Formula: $IR = (N/EH) \times 200,000$

IR = Incidence Rate

N = Total number of occupational injuries and illnesses

EH = Total hours worked by all private industry employees during the calendar year

200,000 = Base for 100 equivalent full-time workers - 40 hours per and 50 weeks per year

Kansas Occupational Injury and Illness Incidence Rates: The division collects data on the entire population of workplace injuries and illnesses in the state of Kansas through its first report of injury form, and stores it in its relational database. Every employer covered under the Workers Compensation Act that has workplace injuries must submit first reports of injury. The severity of each occupational accident or illness and the industrial classification code are mandatory data elements that must be reported by employers to the state. The severity of each accident or illness is exhibited by the numerical code representing the following severity categories: 0-No time lost, 1-Time lost and 2- Fatality. The division's analysts utilized the BLS statistical formula (see above) to calculate the incidence of injury for each severity classification for Kansas's non-federal employment hours for the past ten fiscal years. Data used in the calculation of incidence rates was obtained from the Kansas Labor Market Information Services and division databases.

Kansas Closed Claims Study (CCS) Methodology

The following is a description of the methodology used by the Technology and Statistics Section of the division for the 2004 Closed Claims Study (CCS).

Sample Design: The division consulted with a professor of statistics from Washburn University in order to achieve both efficiency and effectiveness in the CCS Study. Rather than collecting data from the entire population of claims for a calendar year, which would be impractical (as it would result in very large data sets), extremely expensive and labor intensive, the division's researchers randomly sample from the population and make valid inferences about its characteristics using reliable and credible statistical techniques.

Typically, it is preferable to use simple random sampling in a study of this nature. The goal of simple random sampling designs is to ensure that each element in the population has an equal chance of being selected for the study. However, this type of one-stage sampling of carriers is not an appropriate option for the CCS study because the division must have a sample that is sufficiently large and accurately representative of the population in order to perform relevant statistical inference in this way. The sample must also preserve the power of equal probability associated with simple random sampling for statistical purposes. This enables the researchers to process the statistics without having to weight any of the data, thus making the calculations simpler and easier for the public to understand. The paid loss claims are not evenly distributed within the carrier population. The carriers with higher paid losses tend to have a higher proportion of paid loss claims. Therefore, the division needed to sample a larger percentage of these carriers in order to ensure that the sample was representative of the total population. In order to accommodate this situation, the division utilizes a two-stage type of probability sampling procedure known as "disproportionate stratified sampling." Unlike a simple random design, the stratified sample design ensures that different groups within the population will be adequately represented in the sample, thus increasing the accuracy of the parameter estimations. The general strategy employed is to first create strata (subsets of the total population) that are more homogeneous than the population as a whole, and then to sample a different fraction of each stratum population. Then, when combined, the resulting sample will be reasonably representative of the more heterogeneous total population. Furthermore, each carrier is required to sample claims from their database using simple random sampling techniques. This resulting sample will preserve the principle of simple random selection as each carrier of the sample is randomly selected from within each stratum.

The study team estimated that in order to obtain a statistically significant sample, it would be sufficient to sample approximately 35-40 insurance carriers, pools and self-insured organizations. The sampling method is as follows: The population was first stratified according to paid losses. The specific variable used to stratify this population was a percent of total paid losses for all workers compensation claims in the state of Kansas. Subsequent to stratification, the division selected carriers from each stratum by utilizing a random number generator. All carriers in any particular stratum had the same chances of being selected as any other carrier in the same stratum. The selected carriers were then asked to randomly select claims from their own databases that met the CCS study criteria. In consultation with the division's statistician, the

division stratified the population as follows: stratum 1 contains all members of the population with greater than or equal to two percent of the total paid losses; stratum 2 contains all members of the population with greater than or equal to one percent (but less than two percent) of the total paid losses; stratum 3 contains all members of the population with greater than or equal to 0.5 percent (but less than one percent) of total paid losses; stratum 4 contains those members with greater than or equal to 0.25 percent (but less than 0.5 percent) of total paid losses; and stratum 5 contains members with greater than or equal to 0.1 percent (but less than 0.25 percent) of total paid losses. As discussed above, carriers from the strata needed to be sampled disproportionately in order to maintain the principle of simple random sampling. After running a sampling procedure and modifying sampling numbers, final sampling percentages were established for the various strata. One hundred percent of the elements of stratum 1 were selected, 53.8 percent of the elements of stratum 2 were selected, 34 percent of the elements of stratum 3 were selected, 12.9 percent of the elements of stratum 4 were selected, and 11.5 percent of the elements of stratum 5 were selected. This resulted in a total sample of 38 elements from the total population, which satisfies the requirement of 35-40 elements necessary for a statistically significant sample in this case. On occasion, if the situation arises, the director of the division may need to enlarge the study or substitute carriers for political, administrative or financial reasons. In order to accommodate the possibility of this occurrence, the division created a backup list of carriers by first removing the initially selected carriers from the original database and sampling from the resulting list.

Data Collection: The organizations included in the study were then asked by the division to randomly sample from their databases approximately 200 (or less, if they did not have 200) claims for the specified calendar year. The sample was to be taken from each entity's pool of claims, including both medical and indemnity payments. Each claim in the sample also was required to have been open at least one day during the period of January 1, 2003, to December 31, 2003. The expected sample size of both the open and closed claims was approximately 3,765 random claims.

The division secured permission from the National Council on Compensation Insurance to print and use the Detailed Claim Information (DCI) survey instrument in order to create data definitions and structure for the Kansas Closed Claims Study survey. DCI is a national standard for reporting comprehensive claim data from insurance carriers. Adhering to the DCI structure, programmers in the Kansas Department of Labor created two software packages to assist reporting entities. One package was a manual entry system; the other application allowed a text file to be imported electronically. Both products included editing limitations on inputs to certain data fields. In addition, the division's analysts performed data scrubbing on the data sets to ensure that accurate aggregate statistics were reported to the legislature.

To assist reporting organizations in understanding data requirements and use of the new software, the division offered assistance upon request. Following distribution of both software products, reporting entities were asked to provide data by May 31, 2004, on the complete history of their sample of claims. In addition, the division asked that information on charges paid for certain types of services were reported quarterly for claims remaining open within the sample group.

Response Rate: Non-response bias is always a threat to the accuracy of a sample because non-respondents may differ significantly from survey respondents. Typically, in any study such as this one, certain organizations do not respond due to various circumstances, including, but not limited to, bankruptcy, refusal to answer or lost forms. In order to assess whether the CCS sample suffered from non-response bias, the division calculated an estimated response rate for the study. In order to do this, the division first assumed that claims reported by the carriers who had less than the required 200 were, in fact, all of their claims for the year 2003. For the purposes of calculating the non-response rate, the total number of claims not reported was estimated by comparing the number of submitted claims from carriers in the study within the same stratum. Specifically, the statistical mean of the number of claims received per carrier in each stratum was used to estimate the number of claims expected from each non-respondent in the same stratum. This estimation is known as the “mean imputation.” The total number of claim records that the division received was 2,708 out of an estimated 2,860 sampling units. The division used the following standard formula to calculate the response rate, R : $R=1-[(n-r)/n]$, where n = sample size, and r = number of actual responses. The division’s analysts calculated a response rate of approximately 97.12 percent for this study, and thus a corresponding non-response rate of approximately 2.88 percent. The division, in consultation with its statistician, concluded that the sample did not suffer from systematic non-response bias.